

Scene-Based Wave-Front Sensing for Remote Imaging

L. A. Poyneer, K. La Fortune, C. Chan

This article was submitted to *Advanced Wavefront Control*,
San Diego, California,
08/06/2003 – 08/07/2003

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

July 30, 2003

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

Scene-based wave-front sensing for remote imaging

Lisa A. Poyneer^a, Kai La Fortune^a and Carri Chan^{a,b}

^aLawrence Livermore National Lab, Livermore, CA, USA

^bMassachusetts Institute of Technology, Cambridge, MA, USA

ABSTRACT

Scene-based wave-front sensing (SBWFS) is a technique that allows an arbitrary scene to be used for wave-front sensing with adaptive optics (AO) instead of the normal point source. This makes AO feasible in a wide range of interesting scenarios. This paper first presents the basic concepts and properties of SBWFS. Then it discusses that application of this technique with AO to remote imaging. For the specific case of correction of a lightweight optic. End-to-end simulation results establish that in this case, SBWFS can perform as well as point-source AO. Design considerations such as noise propagation, number of subapertures and tracking changing image content are analyzed.

Keywords: adaptive optics, remote imaging, wave-front sensing

1. INTRODUCTION

Adaptive optics (AO) is a technique that corrects in real-time for phase aberrations in an optical system. AO works successfully in many areas, including astronomical telescopes,¹ solar telescopes^{2,3} and vision science. With the exception of solar AO, all these systems use a point-source. There are many interesting cases where the use of AO is desirable, but no point source is available. These include remote imaging from space and along short horizontal or slant paths from the ground. This paper focuses on the case where AO can be used to correct for time-varying aberrations due to a light-weight primary optic. Instead of creating a point source (akin to using a laser guide star) the observed scene can be used to do wave-front sensing. This technique is called scene-based wave-front sensing (SBWFS).

This paper is divided into three parts. First, the SBWFS technique is summarized (a detailed analysis can be found elsewhere⁴) with an emphasis on remote imaging applications. Next, a detailed model of a light-weight optics is presented. This model is used in end-to-end AO simulations to study SBWFS performance. Finally, design considerations and system analysis is conducted.

2. THE SCENE-BASED WAVE-FRONT SENSING ALGORITHM

An AO system using SBWFS is very similar in design to one using a Shack-Hartmann WFS with a point source. Instead of forming an image of the point source on a small area of the WFS CCD, an image is formed instead. A field stop is necessary to control field size and prevent overlap on the WFS CCD. Using these subimages, the average phase gradient is measured in each subaperture across the pupil. If the phase aberration is conjugate to the pupil, all parts of the subimage have the same point-spread-function (PSF). This means that just as the spot formed by a point source is shifted by the average phase gradient,⁵ so is the scene subimage. We assume that each subimage has a small number of pixels (from 8 to 32, preferentially a power of 2) and is Nyquist sampled.

Given these shifted subimages, it is a signal processing problem as how to best estimate the shifts of each image. The best technique for determining this estimate is periodic correlation with use of FFTs. The correlation between two subimages is calculated, and the peak location is estimated with parabolic interpolation. This allows for the necessary subpixel resolution for shifts. In the general case the reference for correlation is a fixed subaperture subimage from the same temporal frame. This reference will change with each time step to ensure the best possible match in image content between subimage and reference.

E-mail: poyneer1@llnl.gov, Telephone: 1 925 423 3360

2.1. The algorithm

Slope estimation involves computing the cross-correlation of two subaperture images (the reference $r[m, n]$ and a subimage $s[m, n]$), finding the maximum and then using that maximum value and the two neighboring values each to determine the estimate of the shifts x_0 and y_0 via parabolic interpolation. The random vector $C[m, n]$ represents the cross-correlation function of these two images. Specifically, this periodic correlation

$$C[m, n] = \sum_i \sum_j r[i - m, j - n] s[i, j] \quad (1)$$

can be computed with FFTs. The maximum of the correlation will be (for whole-pixel shifts) at exactly $C[x_0, y_0]$. For subpixel shifts, the maximum of this correlation function is at $[\Delta_x, \Delta_y]$. We will assume that the maximum of $C[m, n]$ is within half a pixel of the actual shift. For a single estimate of the x-slope, parabolic interpolation is used. This requires the discrete maximum $C[\Delta_x, \Delta_y]$ (as opposed to the true maximum of the continuous correlation) and the two points bracketing it, $C[\Delta_x - 1, \Delta_y]$ and $C[\Delta_x + 1, \Delta_y]$. For notational simplicity, these three points will be referred to as C_0 for the maximum and C_{-1} and C_1 for the neighbors. The estimate of the shift is then:

$$\hat{x}_0 = \Delta_x + \frac{0.5(C_{-1} - C_1)}{C_{-1} + C_1 - 2C_0}. \quad (2)$$

This expression is very difficult to analyze. Because it involves division of random variables, full knowledge of the probability distributions of C_{-1}, C_0, C_1 is required to characterize the resulting random variable. Using a linearization with partial derivatives, the estimate is approximated as

$$\begin{aligned} \hat{x}_0 \approx & \Delta_x + [C_{-1}(m_1 - m_0) + C_0(m_{-1} - m_1) + C_1(m_0 - m_{-1}) \\ & + 0.5(m_{-1} - m_1)(m_{-1} + m_1 - 2m_0)] \times (m_{-1} + m_1 - 2m_0)^{-2}. \end{aligned} \quad (3)$$

Since this is a linear combination of random variables, the mean and variance can be determined with knowledge of only the means and variances of its components. The means, variances and covariances (e.g. m_0, σ_1^2) that appear in the above equations can be easily calculated from the statistical models of the images.⁴ Monte Carlo simulations can also produce estimates of image performance. These agree very closely with analytic predictions. When error standard deviation is above 0.2 pixels, the analytic results (due to the above approximation) begin to underestimate the true error.

2.2. Performance in the zero-shift case

A special case worth considering is when the actual shift between the two images is zero. In a closed-loop system, the image shift will be driven towards null. This simplification also allows easier analysis of slope estimation behavior as illumination conditions change. In this zero-shift case, the two subimages have identical distributions. Therefore $m_{-1} = m_1$ and $\sigma_{-1}^2 = \sigma_1^2$. This reduces the approximation of the estimate [Eqn. 3] to be

$$\hat{x}_0 \approx \Delta_x + \frac{C_{-1} - C_1}{4(m_1 - m_0)} \quad (4)$$

In this special case, the correlation is actually an auto-correlation, so the peak will be at 0 and the means and variances of C_{-1} and C_1 will be equal. Therefore the estimate is unbiased. The variance is

$$\sigma_x^2 = \frac{\sigma_1^2 - \sigma_{-1,1}^2}{8(m_0 - m_1)^2}. \quad (5)$$

The most important term in this equation is the denominator term $(m_0 - m_1)$. As described above, m_0 is the expected value of the maximum of the correlation function (C_0) and m_1 is the expected value one pixel to the side (C_1). The $(m_0 - m_1)$ term is then a measure of the sharpness of the correlation peak. The sharper this peak, the lower the error variance. The correlation function is paired with its Fourier transform partner: the power spectral density. The more impulse-like the correlation function (hence the sharper the peak) the broader the frequency content of the image. This is consistent with the notion that images with more high-frequency

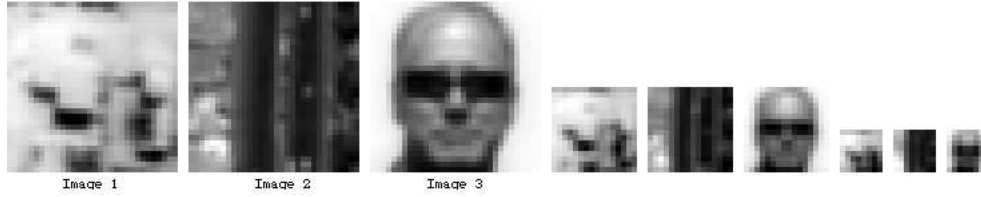


Figure 1. Three sample scenes shown at different resolutions and numbers of pixels. From left to right are 32, 16, and 8 pixel-on-a-side Nyquist-sampled subimages, with the diffraction limit doubling at each step.

content perform better. In the zero-shift case, all that is needed to calculate the error variance of the estimate is knowledge of the subimage statistics. These can be quickly calculated. When done for a wide range of images, the predicted slope estimate variance σ_x^2 reveals substantial variation in image quality. Three sample images (at different resolutions) are shown in Fig. 1.

For each image three different Nyquist-sampled versions are shown, for increasing diffraction limit. This is equivalent to maintaining the same field of view and WFS sampling relative to the diffraction limit while increasing the number of subapertures across the pupil. These three versions are 32, 16 and 8 pixels on a side. For a given size, the performance of these three scenes is quite different. The rms estimation errors σ_x and σ_y are calculated based on the above formula for the zero-shift case. We assigned each image a dynamic range from a minimum of 0 counts received to a maximum of 500 counts received per pixel. For Image 1, $\sigma_x = 0.013$ and $\sigma_y = 0.011$ pixels. For Image 2 $\sigma_x = 0.009$ but $\sigma_y = 0.019$ pixels. The image is much worse in the y-direction. The reason for this is clearly visible in the image content: the road runs along the y-direction. A shift of the image along the road matches well with the reference, because the road is self-similar. Image 3 has $\sigma_x = 0.013$ and $\sigma_y = 0.008$.

None of the analysis above explicitly took in to account the total number of pixels. The performance of a scene is dependent on the frequency content, meaning that a very important design factor is the feature size in the subimage given the system's field of view and resolution. Image 2 is also a good example of this. In the 32×32 version small features (which are cars) are clearly visible. By the time the image is down to the smallest size, σ_y has degraded to 0.087 pixels, even in this well-illuminated case. That's 4.5 times worse than the image at the higher resolution.

The above calculations can be done with a single copy of a subimage in a real system, provided that the level of noise is not too high. This means that on-the-fly estimations of estimation error can be done. An AO system could scan a larger region for the best possible scene, making the system more robust.

2.3. Dealing with changing amounts of light

Performance of the image depends not only on scene content but on the amount and type of illumination. There are two major areas of concern. The first is the total amount of light received, which is primarily a function of the system optical design and the AO control rate. The second is excessive amount of background scatter, which is due to long paths through the atmosphere and can be exacerbated by low visibility. Performance in both of these cases can be predicted analytically.

If the maximum pixel count value is parameterized out with factor f (i.e. $f\tilde{\lambda}[m, n]$), then the formula for error standard deviation becomes

$$\sigma_x(f) = \frac{1}{\sqrt{f}} \frac{(\tilde{\sigma}_1^2 - \tilde{m}_1 - \tilde{\sigma}_{-1,1}^2)^{1/2}}{2\sqrt{2}(\tilde{m}_0 - \tilde{m}_1)}, \quad (6)$$

which is valid for the case where $(f-1)f^{-1} \approx 1$. In this case the signal-to-noise ratio (SNR) is simply proportional to \sqrt{f} , where f is the maximum amount of light per pixel. The standard deviation of the estimate follows the same inverse power law to the SNR as quad-cell centroiding with a point-source does.⁶ Though the constant of the relationship may be different (and is image dependent), this method of wave-front sensing is statistically equivalent to the traditional approach using a point source and centroiding.

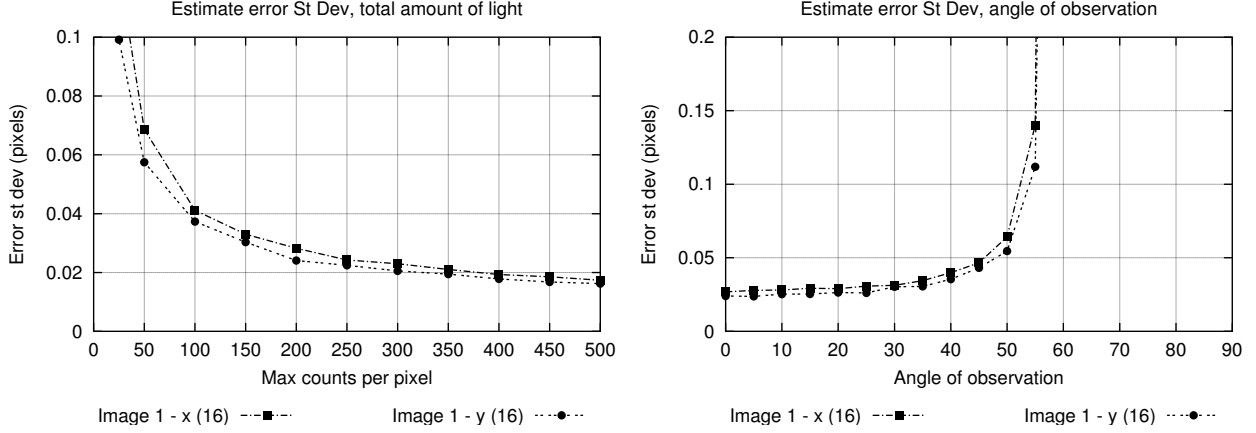


Figure 2. Estimate error standard deviation versus illumination profile, Monte Carlo simulation. [Left]: Changing number of counts received. This curve follows an inverse power law with maximum number of counts. [Right]: Changing angle of observation down through the atmosphere. As the angle increases off normal, the total light decreases and the amount of light in the background dramatically increases.

The case of background illumination can be modeled with a constant parameter b added to the image profile (i.e. $b + f\tilde{\lambda}[m, n]$). The estimate error standard deviation in the zero-shift case is

$$\sigma_x(b, f) = \frac{\{Nb^2 + 2bf^2(\tilde{m}_0 - \tilde{m}_2 + \tilde{t}f^{-1}) + f^3[\tilde{\sigma}_1^2 - (f-1)f^{-1}\tilde{m}_1 - \tilde{\sigma}_{-1,1}^2]\}^{1/2}}{2\sqrt{2}f^2(\tilde{m}_0 - \tilde{m}_1)}, \quad (7)$$

where N is the total number of pixels in a single subimage and t is another image statistic. It should be noted that, unlike the simpler case of a scaling of total illumination, the error standard deviation depends now on the number of pixels in the subimage. Larger images should have worse performance, image content being equal. Smaller images will do better. The sharpness of the correlation peak is still the dominant term in this expression. For both of these cases the expectation of the estimate is independent of the illumination profile. SBWFS is unbiased with respect to background levels, which simplifies the processing of images from the WFS camera.

Analysis was performed to obtain realistic estimates of the amount and type of illumination when observing through the atmosphere. We used a radiometric model based on the radiative transfer equation with some simplifying assumptions including single scattering and decoupled downward and upward irradiances. Using specific parameters such as path length, angle and visibility, a code was written to generate estimates for the level of light received from the image and from the background. Light-level estimates were generated for the case of observing the Earth from space. Results for the two scenarios of changing exposure time and changing angle of observation off normal to the surface are shown in Fig. 2. For the exposure time case, the number of counts is assumed to vary linearly with frame rate. This is the inverse-power-law in SNR as described above. For a variety of scenes performance is very good down to less than 50 counts per pixel. For the angle of observation case, as the image is seen off-normal through thicker layers of atmosphere the total amount of light received decreases while the portion of background light increases. Observation at normal has a maximum of 363 total counts and a background level of 67 counts. In this specific case, once the amount of background light is two-thirds of the total light received (which occurs in this model near 55 degrees), performance begins to degrade rapidly. In the general case this cut-off point will depend on the total amount of light received. Longer exposures are more robust to high background levels than short ones.

2.4. Dealing with larger shifts

Performance can be explicitly analyzed for shifts off-null. Doing so requires direct knowledge or estimation of the image profile at a given shift. As a general trend across images, both the slope estimation error and the error standard deviation increase as the actual shift increases from zero pixels. The practical impact of this is

that a given image may have a higher noise level than indicated by the calculation of σ_x from a single copy of the image. This depends on how large the shifts actually are in closed loop. The estimate error also gradually increases with true shift amount, and peaks around a shift of 0.5 pixels. In closed loop, when shifts are small, this results in a gain change in the system. The slope estimates will may be, for example, only 90% of the correct answer. This should result in increased residual error in the system, but this can be ameliorated by increasing the gain in the control loop.

3. LIGHT-WEIGHT OPTIC AO SIMULATIONS

When the phase aberration is conjugate to the pupil, the above derivation of scene motion based on average phase gradient is completely valid. Because the aberrations are at the pupil, the entire field of view of each subaperture is isoplanatic, i.e. the PSF for each part of the image is the same. In this case each subimage is exactly shifted the way a point-source would be. This is independent of whether or not the object is at infinity (plane wave propagation assumed) or near the aperture (spherical wave propagation assumed). This isoplanaticity also holds for image formation at the science camera. Because light from all parts of the image has the same PSF, the correction of the image field will be uniform.

This scenario is valid in the case of a space-based imaging system with a light-weight optic. The light-weight optic will cause time-varying phase aberrations in the system. We have conducted complete end-to-end simulations of the space-based case. The model for the dynamic phase aberration is discussed first, followed by the end-to-end AO simulation results.

3.1. Light-weight optic phase aberration model

The most practical design for a space-based imaging system would almost certainly be built on reflective optics but designs including diffractive elements are also currently under consideration. The single most massive element in such a system, and hence most desirable to replace, would be the primary optic of the imaging telescope. Designs and prototypes for lightweight reflective optics to replace a conventional primary are already being developed and characterized. T.W. Barbee, Jr. *et. al.*⁷ from LLNL are developing metal nanolaminate thin-shell mirrors. R.C. Romero *et. al.*⁸ from CMA, Inc. are developing carbon fiber reinforced polymer (CFRP) membranes. Other work has been done at Boeing-Rocketdyne and Air Force Research Lab⁹ and at the University of Arizona¹⁰ to develop conventional polymer membranes. All of these approaches have conquered the significant hurdle of obtaining a high optical quality at high spatial frequencies but as the flexibility of a membrane implies, the long range figure of a mirror made out of such a material would be highly unstable.

There are three primary sources of aberrations in a membrane-type optical surface. The simplest source is thermal expansion and contraction of the membrane itself, which results in primarily a defocus term. Because of the elementary nature of this aberration and because it can be removed by a simple repositioning of the optics (not requiring a fully-fledged deformable mirror for correction), it will not be discussed further here. The second most elementary form of aberrations in the system are those that fall into the category of "figuring error". These are aberrations that result from errors in the fabrication process be they inhomogeneities in the membrane material, imperfections in the mandrel, stresses introduced into the membrane when removing it from the mandrel or from some other unidentified source. The nature of these aberrations are as of yet not well known, partially because fabrication techniques have not been stabilized but also because the reluctance of manufacturers to admit any shortcomings in their techniques. The emphasis of recent work has been to prove the ability to reduce high spatial frequency surface errors that would scatter light. The aberrations from figuring errors would be of much lower spatial frequency and would also be static. For this second reason, figuring errors will also not be discussed further because static aberrations are relatively easy to correct.

The third, and most complex, source of aberrations are those introduced by mechanical forces acting on the membrane. The dynamic mechanical aberrations of a membrane have been known and well-understood for years. They are simply linear combinations of Bessel functions. The particular combination depends only on the nature of the excitation of the membrane. We consider a general, and also an inherently worst-case, scenario: a simultaneous excitation of a linear combination of many modes. The amplitude of the modes was assumed to decay exponentially with frequency. This is typical behavior as the amplitude of an aberration is limited by a

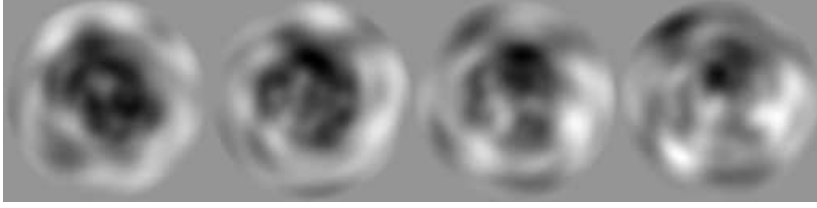


Figure 3. Four realizations of the light-weight optic phase aberration model.

membrane’s ability to stretch. If a mode has more zero-crossings, it follows that its maximum excursion must be smaller. The effective cutoff is adjustable and is set to be the mode where the normalized amplitude is e^{-1} .

Each mode is given a random initial phase and, if there is a theta dependence to the mode, a random initial angle. Note that initially after any excitation, any set of modes that is excited will be correlated in phase and angle. They will, however, decorrelate over time. The purpose is not to model transients in the system that may interrupt imaging momentarily but to model persistent, albeit dynamic, aberrations. As such, the amplitude of the aberrations in the current model do not have a time-dependence. This again describes a worst-case scenario in which the aberrations continue for a long time relative to the desired imaging duration. It is currently unknown exactly how long a membrane will continue to vibrate in a vacuum. But, to be sure, it will be much longer than it would otherwise be when it’s motion is damped by an ambient atmosphere.

Based on the above model, phase screens representing the aberration were generated. A sample set of four independent realizations of the phase aberration is shown in Fig. 3.

3.2. Closed-loop simulation results

The end-to-end AO simulation models the light-weight optic case, where the phase aberrations are pupil-conjugate. The aberrations are time-varying and were generated based on the model above. The harmonic cutoff is the 10th harmonic. The temporal decorrelation time is one-tenth of the system frame rate. For the simulation, the phase has 8×8 samples in each subaperture. There are 16 subapertures across the diameter of the pupil. Each subimage on the WFS camera is Nyquist sampled, giving the WFS one-sixteenth the spatial resolution of the science camera. The field of view of the WFS is one-fourth the width of the science camera. Photoelectron noise on the WFS is generated with poisson random variables. Slope estimation is then done with SBWFS. Next, the phase is reconstructed with the Fourier transform reconstruction¹¹ and the correction is applied with continuous-facesheet DM model. For closed-loop control the correction is calculated with an integrator weight on the previous estimate of 0.99 and a gain on the current phase estimate. The correction is applied with one time-step delay. For comparison, the system can be run noise-free with a point source and centroiding on 4×4 pixel subapertures to provide a measure of best-case performance. Because the edges of the optics are assumed to be fixed, tip-tilt control is not done.

As mentioned above, the use of scenes can result in a gain change in the system. This is clearly evidenced by experiments with the AO simulation code. With a fixed system gain, the ideal point-source model corrected the dynamic phase to a lower level of residual error than some scenes were capable of. Simply increasing the gain enabled the scenes to correct just as well. This indicates that a SBWFS AO-system may need some form of dynamic control over the control loop gain to optimize performance.

Given gain adjustment, a reasonable scene can correct the phase to levels equivalent to that done with a point-source WFS. Fig. 4 shows slices from steady state correction with Image 1 as the scene. The residuals from both the point-source OA and the SBWFS Ao are extremely close. This residual exists due to temporal and fitting errors.

Given a specific scene (Image 1 again) the illumination can be changed. Based on a standard analysis of system error terms,¹ a simplified model is constructed. The mean-square wave-front error σ_ϕ^2 is simply the sum of two independent terms: σ_s^2 , which captures temporal, fitting and all other error terms in the noise-free

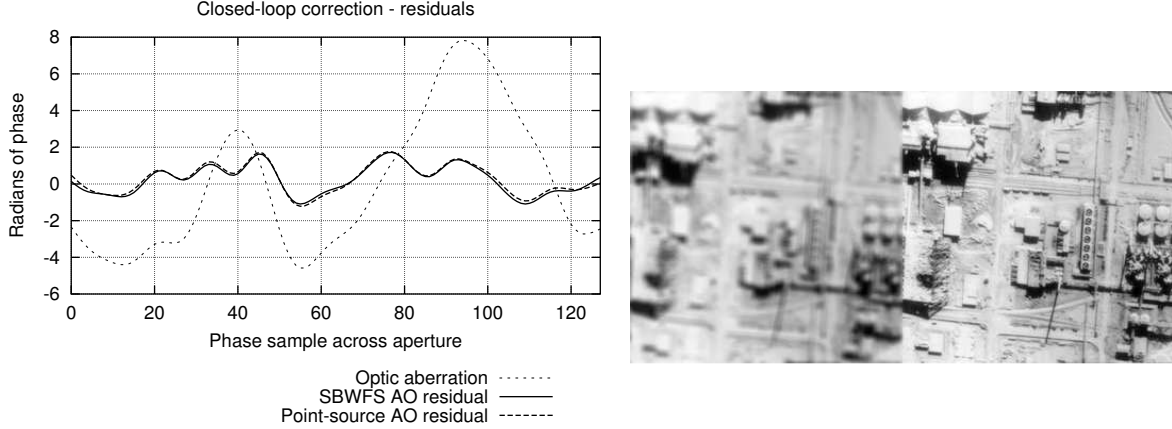


Figure 4. [Left]: Slice of the phase aberration and the residuals at timestep 49 in closed-loop simulation. [Right]: Target image, before and after SBWFS AO correction.

simulation, and σ_w^2 , which captures the propagation of WFS noise. Since an image can have different amounts of noise due to the x- and y-slope estimates, the WFS noise is given by

$$\sigma_w^2 = \frac{\pi^2}{4} N_p \left(\frac{\sigma_x^2 + \sigma_y^2}{2} \right), \quad (8)$$

where the noise propagator N_p is the number that converts WFS noise to phase residual variance. In our simulation the conversion from slope in radians per subaperture to shift in pixels is a factor of π^{-1} . The control loop gain is 0.5. For this simulation $N_p = .667$. Given this relationship, the overall residual error as a function of illumination can be predicted. Fig. 5 shows the total residual mean-squared error versus timestep for four cases with different levels of total light. These cases assume no background and have 50, 25, 10 and 5 counts maximum per pixel per frame. Based on a single copy of the image, σ_x and σ_y were estimated. The right panel of the figure shows the estimated σ_w (based on the above equation) and the average σ_w based on simulation results. Given this particular image and simulation parameters, the AO system is very robust to low light levels. Assuming a closed-loop σ_s^2 of 0.65 squared-radians, and limiting σ_w to 0.35, the RSS value of the slope estimation standard deviations is 1.0 pixel. This would allow adequate closed-loop operation with as few as 10 counts per pixel per frame with no background. This lower limit is dependent on the total error in the rest of the system and the conversion from slope to shift, and will most likely be different for a different system.

4. SYSTEM DESIGN ISSUES

Results presented above show the feasibility of a SBWFS AO system. Several design choices were made above that enabled successful AO operation. These included the number of subapertures, the size of the field of view and the scaling between phase slope and image shift. All of these are necessary to establish during system design. There are some trade-offs involved in these design choices. For example, better sampling of the wave-front mean smaller subapertures and better phase correction, but the amount of light available is decreased, reducing SNR.

The choices of field of view and phase-to-shift scaling are intertwined. Ideally, the field of view will be large enough, given the WFS camera pixel size, to capture images with reasonable feature content. Based on expected phase aberrations and scene imaging properties, the expected amount of shift in both open and closed loop can be predicted. Ideally this shift will not be more than a few pixels in open loop and not more than a few tenths of a pixel in closed loop. This will assure best scene performance, as error standard deviation tends to increase for shifts of null.

Fundamental to the ability of the AO system to work well, regardless of the scene content, is the number of subapertures across the pupil. If the phase aberration's power spectral density is known, the subaperture size

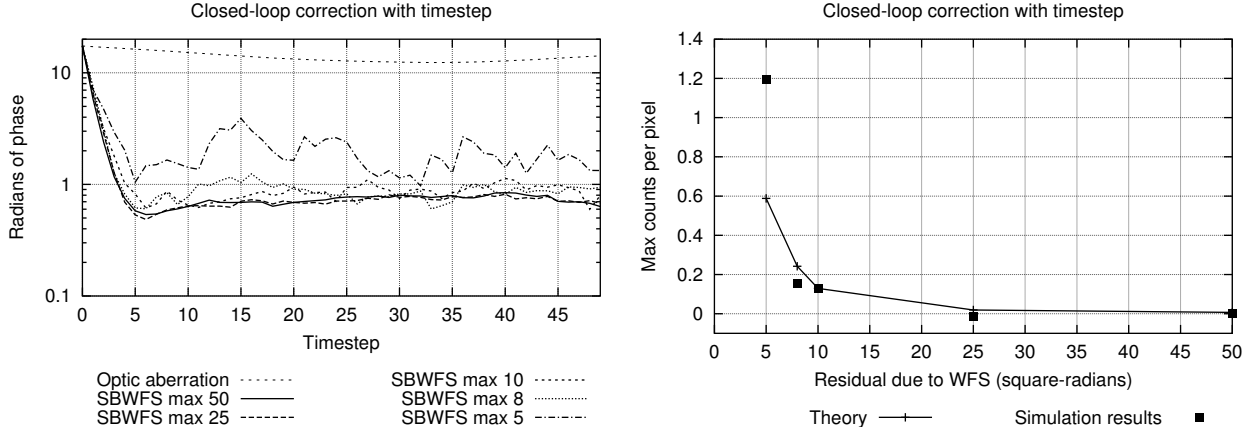


Figure 5. [Left]: Aberration and residuals with time step. As the light level is reduced, the residual becomes both higher and less stable. [Right]: Predictions of σ_w^2 compared with estimations based on trials with Image 1.

should be set such that aliasing is minimized and the phase aberration can be well-corrected given the sampling spatial frequency cutoff $1/2d$. This kind of analysis would only show how well the phase could be corrected given perfect phase measurements.

In reality, slopes are estimated from the subimages. If the phase aberration at each subaperture is dominated by tip and tilt, the image structure will be maintained and the images simply shifted. If significant higher-order phase aberrations exist (primarily if the subapertures are too large) then the image structure will also be distorted, leading to poorer performance in comparing scenes and estimating slopes.

For this latter approach, atmospheric turbulence analysis provides a good place to start: determining the amount of power in tip and tilt in a subaperture. For atmospheric turbulence it is established that 87% of the piston-removed power in an aperture is in tip and tilt.¹² The percent-power metric is equivalent to mean-squared-error of fitting the phase aberration to a linear phase function.

This analysis was conducted to determine the number of subapertures necessary for a given optic aberration profile. For a given phase profile and number of subapertures, Monte Carlo simulations were carried out. For a fixed subaperture location, thousands of random realizations of the optic phase were computed from the phase profile at adequate sampling. Three phase profiles were used. In all cases the amplitude of the higher-order terms (harmonics) falls off exponentially. The harmonic with normalized amplitude e^{-1} was set to be the 5th, 10th and 12th harmonics for these trials. For each, the piston was removed and the percent of energy in tip and tilt was calculated. Based on histograms, the fraction of the total number of trials with at least 85% of the energy in tip and tilt was produced. These results are shown in Fig. 6. For our simulation case above, with sixteen subapertures and the tenth harmonic, the figure shows that 80% of the time in monte carlo simulations the subaperture phase was very-well approximated by tip tilt. For the eight subaperture case, this was true only 40% of the time. The simulation code confirms this analysis. For the eight subaperture case in open-loop, some subapertures see subimage degradation in addition to shifts.

What remains to be determined is a quantitative answer as to how often this 85% criterion must be met for adequate AO performance. For the 5th harmonic cutoff (as shown in Fig. 6) for 8 or more subapertures across the diameter of the optic, this criterion is met at least 80% of the time. For the 10th harmonic case 8 subapertures is clearly inadequate (only 40% of the time is the criterion met) but the 16-20 subaperture range seems satisfactory.

5. TRACKING CHANGES IN SCENE CONTENT

Slope estimation with SBWFS at any step in time depends only on the subimages for that instant. Because a reference from that timestep is used for comparison, the slope estimates contain no tip or tilt information. In the

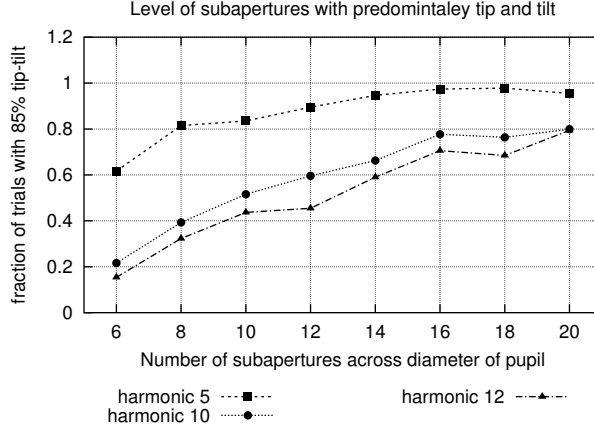


Figure 6. Fraction of trials where at least 85% of the energy in the subaperture phase was in tip and tilt, plotted against number of subapertures across the optic. Results are shown for three harmonic cutoffs - the 5th, 10th and 12th harmonics. As the number of harmonics increases, more subapertures are necessary.

simulation of the light-weight optic the edge was constrained and tip-tilt control was not necessary. However, tip-tilt control will be necessary in the general case if science camera images are taken at exposures longer than the AO frame rate. Just as the shift between two subimages estimates the slope, the shift between two temporal subimages at the same location provides tip-tilt information. This estimate, however, is susceptible to time-varying image structure, such as moving objects or camera motion. To help prevent errors due to changing subimage content, a method to track image motion must be developed.

5.1. Models and methods

Motion detection is achieved by thresholding the normalized variance of each pixel. Previous work has been done using image statistics to isolate moving objects. While these methods are effective for their applications, they either require some knowledge of the image motion which we do not have,¹³ require images over 10 times larger than our resolutions,¹⁴ or require more computation power and memory than we can afford.¹⁵

A pixel's value is modeled as a Poisson process with mean and variance equal to its noise-free intensity value plus Poisson read noise. The mean-normalized variance provides more useful information for motion detection because pixels of higher intensity have higher variance, but equal normalized variance, while pixels with varying intensity have higher normalized variance. In order to differentiate truly moving pixels and not let image motion cause edge pixels to fool us, we categorize pixels as stationary, edge, or moving. The ratio of moving to non-moving pixels allows us to determine what type of image is being viewed: a stationary image in which camera and scene are static, a panning image in which the camera moves over the image, or a moving image in which the camera is static but objects, such as cars, are mobile within it. The motion detection threshold is determined based on the probability model and expected normalized variance of each pixel type. An accurate approximation for the expectation of the normalized variance is the expectation of the sample mean divided by the expectation of the sample variance.

The pixel value of a stationary pixel is the sum of Poisson read-noise and the realization of a Poisson process with parameter equal to the static noise-free pixel value. Where n is the number of samples, the expectation of the mean-normalized variance of a stationary pixel is:

$$E \left[\frac{N_n^2}{n} \right] = \frac{n-1}{n} \approx 1 \quad (9)$$

Random subpixel image shifts can cause edge pixels to be incorrectly flagged as moving. An edge is modeled as being one pixel wide. The intensity of an edge pixel is a random variable that shifts from a mean equal to the average of the intensity levels on each side of the edge. Each shift is a zero-mean Gaussian random variable with

variance σ_s^2 on the order of .01 pixels. Where n is the number of samples, λ is the Poisson read-noise parameter, α and β are the noise-free intensity levels on each side of the edge, the expectation of the mean-normalized variance of an edge pixel is:

$$E[N_n^2] = \frac{n-1}{n} \left[\frac{0.25\alpha + 0.25\beta + \lambda + (\alpha - \beta)^2 \sigma_s^2}{0.5\alpha + 0.5\beta + \lambda} \right] \quad (10)$$

For an edge that goes from 0 counts to 250 counts, the expectation of the normalized variance is over 5 times larger than the expectation for a stationary pixel.

A moving pixel begins at some intensity level and increases with a constant slope k , for n samples. Where h is the sum of starting intensity value and λ which is the Poisson read-noise parameter, the expectation of the mean-normalized variance of a moving pixel is:

$$E[N_n^2] = \frac{6(k + 2kh)(n+1)(1 - n^{-1}) + 2k^2(2n+1)(n+1) - 3k^2(n+1)^2 - 12hk(n^2 - 1)n^{-1} + 12h(1 - n^{-1})}{6k(n+1) + 12h} \quad (11)$$

For a moving pixel that moves from 0 counts to 250 counts with a slope of 50, the expectation of the normalized variance is nearly 8 times larger than the expectation for an edge pixel with the same range. This large discrepancy suggests that we should be very successful using the normalized variance for motion detection.

5.2. Results

Simulations which varied in range of peak/base values and number of samples verified the calculated, expected values of the normalized variance.

Assuming subimages between 16×16 and 32×32 pixels, most moving objects, such as cars, are likely to be on the order of 3 to 5 pixels. As such, a reasonable size for n is 5 because a car of 5 pixels takes 5 frames to fully traverse a single pixel. Requiring more frames would be disadvantageous. In order to use SBSHWFS to find slopes, the frames must be correlated, which is unlikely when frames are separated by many time steps. Also, if the AO system must wait for more frames, it might not be able to correct the phase aberrations quickly enough.

The expected normalized variance for edge and moving pixels depends on range and background level. Consequently, a good threshold depends on the illumination and background levels of the image. Too low a threshold will result in many false alarms while too high of one will result in many misses. It is the edge and moving pixels which cause these errors. Using the maximum and minimum of all pixel values received, we calculate the maximum expected normalized variance for any edge pixel within the image. Because the maximum expected normalized variance for moving pixels is over 5 times larger than for edge pixels, the threshold is 1.25 times the maximum expected normalized variance of edge pixels. This minimizes the number of misses and false alarms

By thresholding the normalized variance of individual pixels, motion detection is successfully achieved. For noise-free images, the motion detection is extremely accurate for 32×32 and 16×16 images (figure 5a, 5b). These images are of a car driving down a street. The white with black border isolate the detected pixels of the moving car. The threshold is based on the expected normalized variance of edges which makes it likely that some edge pixels will exceed the threshold. Stationary pixels can also cause false alarms if they have high variances due to the randomness of noise. As the signal to noise ratio decreases, the false alarms in noisy and shifted images increases. Figure 7c shows 8 false alarms that are dispersed randomly. Assuming that most cars are 3 to 5 pixels, it is unlikely that a single stranded pixel with high normalized variance is part of a moving object. We lowpass filter an image mask where moving pixels are equal to 1 and non-moving pixels are equal to 0. Single pixels with high normalized variance are filtered out leaving behind pixels which are most likely to be part of a moving car. The results of this lowpass filter technique can be seen in figure 7d. Despite the noise and image shifts, the moving car is detected just as accurately in 7d as in the original image in 7b.

Illumination and background levels significantly affect the signal to noise ratio and, hence, the performance of motion detection. The percentage of false alarms is defined as the number of pixels falsely marked as moving in that frame divided by the total number of pixels in the frame. The percentage of misses is defined as the

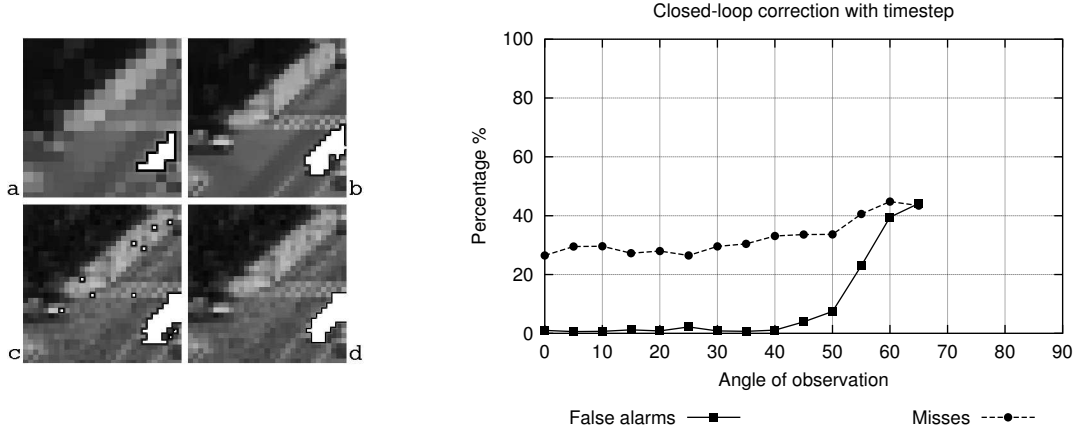


Figure 7. [Left]: Motion detection of a car driving on a street. The white with black border highlight the pixels detected as moving. Top row: a) 16×16 image; b) 32×32 image, both with no noise. Bottom row: c) 32×32 noisy and shifted; d) same image with lowpass filter on results. [Right:] Plot of false alarms and misses versus angle of observation.

number of pixels not marked as moving in the noisy image divided by the number of pixels marked as moving in the noise-free version of that frame. Motion detection performance is very consistent with exposure level, only degrading for fewer than 50 counts per pixel. Figure 7 shows that increasing background level causes a steep increase in false alarms. This occurs at nearly the same angle as loss of σ_x performance as shown in Fig. 2. Beyond this point most scenes have poor quality, so motion detection will not be necessary since the images may be too poor to use.

Thresholding mean-normalized variance to achieve motion detection is more successful when a small fraction of the image is moving and when there are few sharp contrast edges. The number of false alarms decreases significantly for these types of images. This disparity is most likely due to image shifts which cause the variance of a pixel to increase. For pixels that are stationary and are surrounded by mostly stationary pixels, these increases are not large enough to initiate a false alarm. However, edge pixels may shift to cause false alarms and moving pixels may shift to cause pixels previously traversed to be marked. Fortunately, this phenomenon does not hurt our usage of this motion detection algorithm. If images have enough motion to cause so many false alarms, SBSHWFS will be unsuccessful for slope estimation.

Using a threshold on the mean-normalized temporal variance of individual pixels proved to be an efficient and effective method for motion detection. While many standard image-processing techniques failed on our low-resolution images, this method successfully detected moving objects on our 16×16 and 32×32 images.

6. CONCLUSIONS AND FUTURE DIRECTIONS

Scene-based wave-front sensing is a viable method for using AO without a point source. By correlating subimages of the observed scene, slope estimates are calculated. Scene performance can be predicted with a single copy of an image. Performance is robust to short exposures and background, until background becomes dominant in the image. End-to-end simulations with a dynamic light-weight optic phase aberration confirm that SBWFS can work as well as point-source AO. Further analysis is considered to find the best number of subapertures for AO correction. Methods to track a scene content changes with time work well on even small sets of data, allowing good identification of changing scene content for tip-tilt control.

The case of distributed aberration (i.e. horizontal path imaging) is currently under investigation. This problem is more complex, not only due to the standard challenges of measuring and correcting distributed aberrations but due to anisoplanatism within subimages.

ACKNOWLEDGMENTS

Thanks go to Jim Brase, Carmen Carrano and Erik Johansson. This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48. The document number is UCRL JC-?.

REFERENCES

1. J. Hardy, *Adaptive Optics for Astronomical Telescopes*, Oxford University Press, Oxford, 1998.
2. T. R. Rimmelle, "Solar adaptive optics," in *Adaptive Optical Systems Technology*, P. Wizinowich, ed., *Proc. SPIE* **4007**, pp. 218–231, 2000.
3. T. Rimmelle, O. von der Luehe, P. Wiborg, A. Widener, R. Dunn, and G. Spence, "Solar feature correlation tracker," in *Active and adaptive optical systems*, M. Ealey, ed., *Proc. SPIE* **1542**, pp. 186–193, 1991.
4. L. A. Poyneer, "Scene-based wave-front sensing: analysis and simulation," *Applied Optics*, p. in press, 2003.
5. M. C. Roggemann and B. Welsh, *Imaging through Turbulence*, CRC Press, New York, 1996.
6. G. Tyler and D. Fried, "Image-position error associated with a quadrant detector," *JOSA A* **72**, pp. 804–808, 1982.
7. G. S. Hickey, S.-S. Lih, and T. W. Barbee, "Development of nanolaminate thin shell mirrors," in *Highly Innovative Space Telescope Concepts*, H. A. MacEwen, ed., *Proc. SPIE* **4849**, pp. 67–76, 2002.
8. P. Chen, C. Bowers, D. Content, M. Marzouk, and R. Romeo, "Advances in very lightweight composite mirror technology," *Opt. Eng.* **39**, pp. 2320–2329, 2000.
9. J. R. Rotge and S. C. Dass, "Progress toward large-aperture membrane mirrors," in *Imaging Technology and Telescopes*, J. Bilbro and et al, eds., *Proc. SPIE* **4091**, pp. 74–82, 2000.
10. N. J. Woolf and J. W. Bilbro, "Gossamer optics technology workshop," in *Imaging Technology and Telescopes*, J. Bilbro and et al, eds., *Proc. SPIE* **4091**, pp. 90–94, 2000.
11. L. A. Poyneer, M. Troy, B. Macintosh, and D. Gavel, "Experimental validation of Fourier transform wave-front reconstruction at palomar observatory," *Optics Letters* **28**, pp. 798–800, 2003.
12. R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am. (A)* **66**, pp. 207–11, 1976.
13. Y. Ren, C. Chua, and Y. Ho, "Motion detection with non-stationary background," in *Image Analysis and Processing*, *Proc. IEEE* **11**, pp. 78–83, 2001.
14. Haritaoglu, D. Harwood, and L. Davis, "A fast background scene modeling and maintenance for outdoor surveillance," in *Pattern Recognition*, *Proc. IEEE* **4**, pp. 179–183, 2000.
15. T. H. S. Blostein, "Detecting small, moving objects in image sequences using sequential hypothesis testing," *IEEE Trans. Sig. Proc.* **39**, pp. 1611–1629, 1991.